PATENT ABSTRACTS OF JAPAN

(11)Publication number : **10-288996**
(43)Date of publication of application : **27.10.1998**

(51)Int.Cl. **G10L 3/00**
G10L 3/00

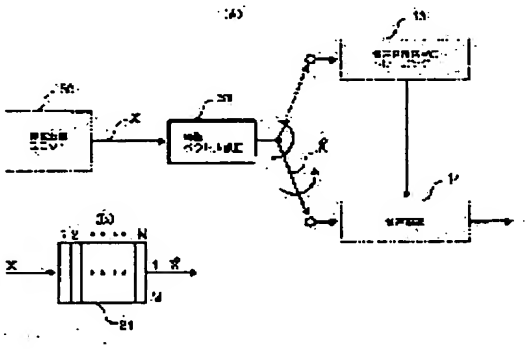
(21)Application number : **10-097547**
(22)Date of filing : **09.04.1998**

(71)Applicant : **NOKIA MOBILE PHONES LTD**
(72)Inventor : **LAURILA KARI**
VIIKKI OLLI

(30)Priority
Priority number : **97 971521** Priority date : **11.04.1997** Priority country : **FI**

(54) SPEECH RECOGNITION METHOD AND SPEECH SIGNAL PROCESSOR

(57)Abstract:
PROBLEM TO BE SOLVED: To correct a feature vector to be decided at a speech recognition time for correcting the effect of a noise by defining the mean value and the standard deviation of the characteristic vector, using these parameters and normalizing the characteristic vector.
SOLUTION: The mean value and the standard deviation of the characteristic vector are defined, and the feature vector is normalized using these parameters. In such a case, the feature vector is normalized using a sliding normarization buffer preferably. In a speech recognition device, a front end 30 forms the feature vector Xi (i=1-M) at a prescribed interval. The feature vector X is stored in a normalization buffer 31, and thus, the mean value and the standard deviation related to respective feature vector components Xi are calculated. Thereafter, the feature vector component Xi to be recognized is normalized using a calculated normalization coefficient in the block 31. In such a case, the normalization buffer 31 that e.g. a length N is fixed is slid on the feature vector X.



LEGAL STATUS
[Date of request for examination]
[Date of sending the examiner's decision of rejection]
[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]
[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

(19)日本国特許庁（ＪＰ）　(12)公開特許公報（Ａ）　(11)特許出願公開番号

特開平10－288996

(43)公開日　平成10年(1998)10月27日

(51)Int.Cl.[6]	識別記号	ＦＩ	
Ｇ１０Ｌ　3/00	５３５	Ｇ１０Ｌ　3/00	５３５
	５２１		５２１Ｆ

審査請求　未請求　請求項の数６　ＯＬ　（全 ９ 頁）

(21)出願番号　特願平10－97547

(22)出願日　平成10年(1998)４月９日

(31)優先権主張番号　９７１５２１

(32)優先日　1997年４月11日

(33)優先権主張国　フィンランド（ＦＩ）

(71)出願人　590005612
ノキア　モービル　フォーンズ　リミティド
フィンランド国，エフアイエヌ－02150　エスポー，ケイララーデンティエ　4

(72)発明者　カリ　ローリラ
フィンランド国，エフアイエヌ－33720　タンペレ，インシンオーリンカトゥ　64　アー　14

(72)発明者　オリ　ピッキー
フィンランド国，エフアイエヌ－33100　タンペレ，ピンニンカトゥ　21　アー　7

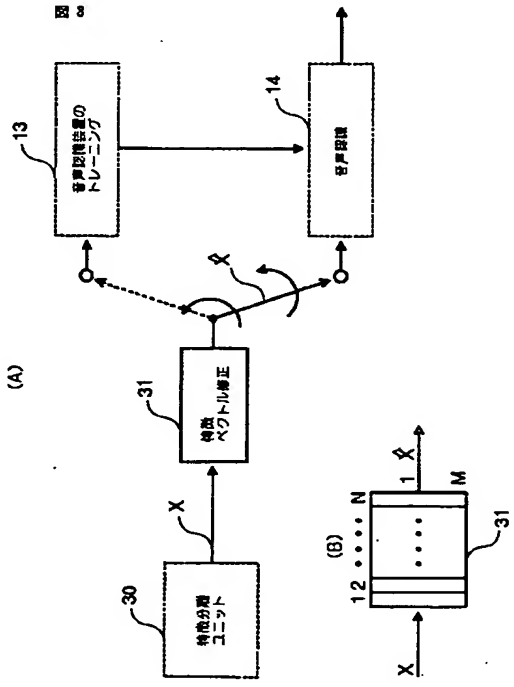
(74)代理人　弁理士　石田　敬　（外４名）

(54)【発明の名称】　音声認識方法及び音声信号処理装置

(57)【要約】

【課題】　改良された音声認識方法及び音声信号処理装置を提供する。

【解決手段】　音声認識装置の分析ユニットで作られた特徴ベクトルが雑音の影響を補償するために修正される。本発明では、スライディング正規化バッファー（３１）を使って特徴ベクトルを正規化する。本発明の方法により、音声認識装置のトレーニング段階が実際の音声認識段階での騒音環境とは異なる騒音環境で実行された場合に、音声認識装置の性能が向上する。



【特許請求の範囲】
【請求項１】 認識可能な音声信号を時間上で特定の長さの連続するフレームに分割し、フレームあたりに少なくとも１つの、該音声信号を説明するパラメータを作るために各音声フレームを分析し、特定のフレームに関連する前記パラメータを記憶し、前記パラメータを修正し、その修正されたパラメータを使って音声認識を実行する音声認識方法であって、連続するパラメータのうちの一部だけを定期的に記憶し、前記の修正されたパラメータを作るために定期的に記憶されるパラメータに基づいて少なくとも１つのパラメータを修正するようになっていることを特徴とする音声認識方法。
【請求項２】 Ｎ個の連続するパラメータに基づいて、次の各量すなわち平均値及び標準偏差のうちの１つを前記修正のために確定するようになっており、Ｎは整数であることを特徴とする請求項１に記載の方法。
【請求項３】 パラメータの前記修正は、前記各量の１つに関連する正規化から成ることを特徴とする請求項２に記載の方法。
【請求項４】 音声信号を時間上で分割して連続するフレームとするための手段（２１）と、音声フレームを分析して該音声信号を説明する少なくとも１つのパラメータを作るための手段（１１，３０）と、該パラメータを記憶するための記憶手段（３１）と、前記パラメータを修正して修正済みパラメータを作るための手段（３１）と、その修正済みパラメータを使って音声を認識するための手段（１４）とから成る音声信号処理装置であって、前記記憶手段（３１）は前記の連続するパラメータのうちの一部だけを定期的に記憶するようになっており、該パラメータを修正するための前記手段（３１）は、前記修正済みパラメータを作るために該記憶手段（３１）に定期的に記憶されたパラメータに基づいて該音声信号を説明する該パラメータを修正するようになっていることを特徴とする音声信号処理装置。
【請求項５】 前記記憶手段（３１）は一定の長さのバッファー（３１）から成ることを特徴とする請求項４に記載の装置。
【請求項６】 前記記憶手段（３１）は長さが変化し得るバッファー（３１）から成ることを特徴とする請求項４に記載の装置。
【発明の詳細な説明】
【０００１】
【発明の属する技術分野】本発明は音声認識に関し、特に音声認識において決定されるべき各特徴ベクトル（feature vector）を修正する方法に関する。本発明は、音声認識を改良するために本発明のこの方法を使用する装置にも関する。
【０００２】
【従来の技術】本発明は自動的音声認識に関し、特に、ヒドゥン・マルコフモデル（Hidden Markov Models (HMM)）に基づく音声認識に関する。ＨＭＭに基づく音声認識は、認識可能な単語の統計的モデルに基づいている。認識段階においては、発音された単語についてマルコフチェーンに基づいて観測結果及び状態遷移が計算されて、音声認識装置のトレーニング段階で記憶された、その発音された単語に対応するモデルが確率に基づいて決定される。例えば、ヒドゥン・マルコフモデルに基づく音声認識方法は下記の参考文献において解説されている：”１９８９年２月のＩＥＥＥ会報第７７巻第２号の中のＬ．ラビナーの”音声認識におけるヒドゥン・マルコフモデルと選択されたアプリケーションについての指導”（"L. Rabiner, "A tutorial on Hidden Markov Models and selected applications in speech recognition", Proceedings of the IEEE, Vol. 77, No.2, February 1989.
【０００３】
【発明が解決しようとする課題】現在の音声認識装置の問題は、騒々しい環境の中では認識精度が著しく低下することである。また、特に音声認識装置が動作するときの騒音条件が該音声認識装置のトレーニング段階での騒音条件と異なる場合には該音声認識装置の性能は低下する。音声認識装置が使用されることのある全ての騒音環境の影響を考慮に入れることは不可能であるので、このことは実際問題としては音声信号認識システムにおいて最も解決しにくい問題の１つである。音声認識装置を利用する装置のユーザにとっての正規の立場は、音声認識装置のトレーニングは通常は殆ど騒音のない環境で行われるけれども、その音声認識装置が例えば車内などの動作環境で使われるときには周囲の交通やその車自体から生じる暗騒音はトレーニング段階での殆ど静穏な暗騒音レベルとは著しく違っているということである。
【０００４】音声認識装置の性能が使用されるマイクロホンに左右されることも現在の音声認識装置の問題である。特に音声認識装置のトレーニング段階で使われるマイクロホンが実際の音声認識段階で使われるマイクロホンとは違っている場合には、その音声認識装置の性能は著しく低下する。
【０００５】特徴ベクトルを計算する際の雑音の影響を消去するために幾つかの方法が開発されている。しかし、それらの方法を利用する音声認識装置は決まったコンピュータ／ワークステーションのアプリケーションでのみ利用可能であり、それらの音声認識装置では音声はオフラインで認識される。それらの方法では、認識されるべき音声はコンピュータのメモリーに記憶されるのが普通である。通常、記憶される音声信号の長さは数秒である。その後、特徴ベクトルは、計算時に、ファイル全体の内容から確定される各パラメータを使って修正される。記憶される音声信号の長さの故に、その種の方法を実時間音声認識に適用することはできない。
【０００６】また、正規化方法も設けられており、その

方法では音声及び雑音の両方が自分の正規化係数を持っていて、それらは音声活性検出器（ＶＡＤ）を用いて適応的に更新される。適応的更新（ａｄａｐｔｉｖｅ ｕｐｄａｔｉｎｇ）であるために、各正規化係数の更新には遅延が伴うので、正規化プロセスは実用上充分な速さでは実行されない。また、この方法もＶＡＤを必要とするけれども、その動作は、信号対雑音比（ＳＮＲ）の値が低い音声認識アプリケーションではしばしば余りに不正確である。この方法も、前記の遅延の故に実時間要件を満たさない。

【０００７】

【課題を解決するための手段】上記の問題を解決する音声認識方法及び装置が発明されており、その方法及び装置により音声認識時に決定される特徴ベクトルは雑音の影響を補償するために修正される。特徴ベクトルの修正は、特徴ベクトルの平均値と標準偏差とを定義し、それらのパラメータを使って特徴ベクトルを正規化することによって実行される。本発明の好ましい実施例では、スライディング正規化バッファー（ｓｌｉｄｉｎｇ ｎｏｒｍａｌｉｓａｔｉｏｎ ｂｕｆｆｅｒ）を使って特徴ベクトルを正規化する。本発明では、特徴ベクトルの正規化パラメータの更新は殆ど遅延無しで実行され、実際の正規化プロセスにおける遅延は充分に小さいので実時間音声認識アプリケーションを実現することができる。

【０００８】また、本発明の方法によれば、音声認識装置の性能を、使用するマイクロホンに左右されにくくすることができる。本発明によれば、音声認識装置の実験段階と認識段階とで異なるマイクロホンが使われる場合にも、同じマイクロホンがトレーニング段階及び認識段階で使われる場合と殆ど同じ程度に、音声認識装置の高い性能が達成される。

【０００９】本発明は、請求項１及び４の特徴付け部分に記載されている事項を特徴とする。

【００１０】

【発明の実施の形態】図１は、本発明に適用できる公知の音声認識装置の構造を示すブロック図である。通常、音声認識装置の動作は、図１に示されているように、主要な２種類の活動、即ち実際の音声認識段階１０－１２，１４－１５と音声トレーニング段階１３とに分けられる。音声認識装置はマイクロホンから入力として音声信号ｓ（ｎ）を受け取り、この信号は、例えば８ｋＨｚのサンプリング周波数及び１サンプルあたり１２ビットの分解能を使用するＡ／Ｄ変換器１０によってデジタル形に変換される。通常、音声認識装置はいわゆるフロント・エンド１１を有し、ここで音声信号が分析されて特徴ベクトル１２がモデル化される。特徴ベクトルは特定の期間中の該音声信号を描写するものである。特徴ベクトルは、例えば１０ｍｓ間隔で確定される。特徴ベクトルを、数種類の手法でモデル化することができる。例えば、特徴ベクトルをモデル化するための数種類の手法が下記の参考文献で解説されている：１９９３年９月のＩＥＥＥ会報第８１巻、第９号、ｐｐ．１２１５－１２４７，に掲載されているＪ．パイコーンの”音声認識における信号モデル化手法（J. Picone, "Signal modelling techniques in speech recognition", IEEE Proceedings, Vol. 81, No. 9, pp. 1215-1247, September 1993.本発明において使用される特徴ベクトルは、いわゆるメルー周波数セプストラル係数（Mel-Frequency Cepstral Coefficients (MFCC)）を確定することによりモデル化される。トレーニング段階で、音声認識装置により使用される単語について音声認識装置のトレーニング・ブロック１３において、特徴ベクトルによってモデルが作成される。モデル・トレーニング１３ａにおいて、認識可能な単語についてモデルが決定される。トレーニング段階において、モデル化されるべき単語の復唱（ｒｅｐｅｔｉｔｉｏｎ）を利用することができる。モデルはメモリー１３ｂに記憶される。音声認識時に、特徴ベクトルは現実の認識装置１４に送られ、この装置は、ブロック１５ａにおいて、トレーニング段階時に構成されたモデルと認識可能な音声から構成されるべき特徴ベクトルとを比較して、認識結果についての判定をブロック１５ｂで行う。認識結果１５は、音声認識装置を使用する人により発音された単語に最もよく対応する、音声認識装置のメモリーに記憶されている単語を表示する。

【００１１】図２は、本発明に適用できるフロント・エンド１１の公知の分析ブロックの構造を示している。通常、フロント・エンド１１は、音声認識に関連する周波数を強調するためのプリエンファシス・フィルター２０を有する。通常、プリエンファシス・フィルター２０は、例えば、Ｈ（ｚ）＝１－０．９５Z^{-1}のレスポンスを有する１次ＦＩＲフィルターなどの高域通過フィルターである。次に、ブロック２１において、フィルタリングされた信号からＮサンプルの長さの各フレームが形成される。例えば、Ｎ＝２４０のサンプル長を使って、８ｋＨｚのサンプリング周波数で３０ｍｓのフレーム構造が作られる。通常、連続するフレーム同士がＳ個の連続するサンプル（例えば１０ｍｓ）の程度に重なり合ういわゆるオーバーラップ手法を使って各音声フレームを形成することもできる。ブロック２３において音声信号について高速フーリエ変換（ＦＦＴ）周波数表示をモデル化する前に、例えば、ブロック２２においてハミングウィンドウ（Ｈａｍｍｉｎｇｗｉｎｄｏｗ）などを使ってスペクトル推定値の精度を向上させるためにいわゆるウィンドウイング（ｗｉｎｄｏｗｉｎｇ）を実行することもできる。次に、信号のＦＦＴ表示をメル・ウィンドウイング・ブロック（Ｍｅｌ ｗｉｎｄｏｗｉｎｇ ｂｌｏｃｋ）２４においてメル周波数表示に変換する。メル周波数表示への変換は、それ自体としては当業者に知られている。メル周波数表示への変換は参考原典”ＩＥＥＥ会報第８１巻、第９号に掲載されているＪ．パイコー

ンの”音声認識における信号モデル化手法（J. Picone, "Signal modelling techniques in speech recognition"）”で解説されている。この周波数変換で、いろいろな周波数に対する耳の非線形の感度を考慮に入れる。通常、使用される周波数帯域の数（ｋ）はｋ＝２４であってよい。実際の特徴ベクトル１２，即ちいわゆるセプストラル係数ｃ（ｉ）は、ブロック２５で形成された２６個の対数メル値に対していわゆる離散余弦変換（ｄｉｓｃｒｅｔｅ ｃｏｓｉｎｅ ｔｒａｎｓｆｏｒｍａｔｉｏｎ）（ＤＣＴ）を実行することによって得られる。この離散余弦変換に例えば次数Ｊ＝２４を使用することができる。通常、ＤＣＴ係数ｃ（ｉ）（ｉは余弦項のインデックスである）のうちの半分だけが使われる。通常、実際の特徴ベクトルは、いわゆる第１段及び第２段の差信号ｄｃ（ｉ）及びｄｄｃ（ｉ）を計算することによって音声の変動過程（ダイナミックス）に関する情報も包含する。ブロック２７においてｄｃ（ｉ）＝ｃ（ｉ）－ｃ（ｉ－１）及びｄｄｃ（ｉ）＝ｄｃ（ｉ）－ｄｃ（ｉ－１）を推定することにより、離散余弦変換ブロックの連続する出力ベクトルからこれらの差信号を決定することができる。これらの２６個の追加のパラメータが考慮される場合には、特徴ベクトルの長さは例えば１３＋２６＝３９パラメータとなる。

【００１２】図３（Ａ）及び（Ｂ）は本発明の第１実施例の音声認識装置の構造を示す。フロント・エンド３０は１０ｍｓ間隔で出力信号として特徴ベクトルX_i，ｉ＝１・・・Ｍ（例えばＭ＝３９）を作成する。特徴ベクトルは正規化バッファー３１に記憶され、これにより各特徴ベクトル成分X_i，ｉ＝１・・・Ｍ、についての平均値μ_i及び標準偏差σ_iが次のように計算される：

【数１】

$$\mu_i = \frac{1}{N}\sum_{t=1}^{N} x_{t,i}, i = 1,\ldots,M \qquad 1$$

【数２】

$$\sigma_i = \sqrt{\frac{1}{N}\sum_{t=1}^{N}(x_{t,i}-\mu_i)^2}, i = 1,\ldots M \qquad 2$$

式（１）及び（２）において、Ｎは正規化バッファー（ｎｏｒｍａｌｉｓａｔｉｏｎ ｂｕｆｆｅｒ）の長さであり、Ｍは特徴ベクトル（ｆｅａｔｕｒｅ ｖｅｃｔｏｒ）の長さである。この後、ブロック３１において、計算された正規化係数μ_i，σ_iを使って、認識されるべき特徴ベクトルの成分X_iが正規化される。正規化され認識されるべき特徴ベクトルＸは、図４に示されているように正規化バッファー３１の中央に置かれる。

【数３】

$$\hat{x}_I = \frac{x_i - \mu_i}{\sigma_i}, i = 1,\ldots,M \qquad 3$$

【００１３】段落番号〔００１３〕から〔００１４〕までに現れる（イ）は、下記表１に示す記号を表す。

【表１】

（イ）	$\hat{x}$

正規化された特徴ベクトル（イ）はそれが音声認識装置のトレーニング段階であるのか実際の音声認識段階であるのかにより、音声認識ユニット１４又はトレーニング・ブロック１３に入力信号として送られる。本発明の第１実施例の方法では、長さ（Ｎ）が固定されている正規化バッファーが使われ、このバッファーは特徴ベクトル上をスライドさせられる。このスライディング正規化バッファーがあるので、この方法を実時間音声認識システムで実行することもできる。正規化バッファー３１はＮ＊Ｍサンプルのサイズのバッファーであり、通常は、デジタル信号処理装置（ＤＳＰ）の内部メモリー構造又は外付けメモリーを使って該ＤＳＰをプログラミングすることによって音声認識ユニットと関連させてこのバッファーを実現することができる。本発明の実施例の解決法においては、正規化バッファーは１００の特徴ベクトルの長さを持っている。一度に正規化され認識されるべき特徴ベクトルは正規化バッファー３１の中央に置かれる。正規化されるべき特徴ベクトルは正規化バッファーの中央に置かれるので、音声認識には正規化バッファーの長さであるＮの遅延が生じる。本例の各パラメータを使うときには、遅延は１００＊１０ｍｓ＝１秒である。しかし、次に説明するように音声認識の始めに該バッファーの長さの一部分だけを使うことによって、この遅延を半分にすることができる。

【００１４】図５及び図６は、フローチャートの形で、本発明の方法の作用を示している。音声認識の始めに、正規化バッファーの全長の半分Ｎ／２が使用されるまで該正規化バッファーは充填される（ブロック１００－１０２）。その後、平均値及び標準偏差の各ベクトルμ_i，σ_iが計算され（ブロック１０３）、はじめのＮ／２個の特徴ベクトルを使って第１特徴ベクトルが正規

化される。ブロック１５ｂ（図１）で公知の手法に従ってビタビ復号（Viterbi decoding）によりこの正規化された特徴ベクトル（イ）に対して実際の音声認識プロセスが実行される。次に、新しい特徴ベクトルが緩衝記憶され（ブロック１０４）、記憶されている（Ｎ／２＋１）個の特徴ベクトルを使って新しい正規化係数が計算され、第２の特徴ベクトルが正規化されて、それに対して認識が実行される（ブロック１０３）。これに対応するプロセスが正規化バッファーが満杯になるまで続けられる。このときフローチャートにおいてブロック１０５からブロック１０６への移行が行われる。このことは、始めのＮ／２個の特徴ベクトルが認識され終わっていて、正規化されるべき特徴ベクトルが正規化バッファーの中央に位置していることを意味する。このとき該バッファーはＦＩＦＯ原理（先入れ先出し）に従ってスライドされて、新しい特徴ベクトルが計算され認識され終わったならば（ブロック１０７）、最も古い特徴ベクトルが正規化バッファーから除去される（ブロック１０６）。認識段階の終わりに（ブロック１０７）、正規化バッファーに記憶されている値を使って正規化係数が計算される。これらの正規化係数が最後のＮ／２個の特徴ベクトルの認識と関連して使用される。平均値及び標準偏差は、正規化されていない特徴ベクトルを使って計算される。Ｎ個の特徴ベクトルの全てに対して音声認識が実行され終わると（ブロック１０８）、音声認識装置は認識可能な単語の結果をモデル化する（ブロック１０９）。

【００１５】本発明の第２の実施例では、正規化バッファーの長さは音声認識中に変化することがある。音声認識開始時には長さが比較的に短い（例えばＮ＝４５）バッファーを使うことができ、例えば各フレーム（３０ｍｓ）について音声認識が進むに連れて、緩衝記憶されるべき信号の長さを大きくしてゆくことができる。この様に、本発明の第１実施例に対する例外として、正規化されるべき特徴ベクトルはバッファーの中央の特徴ベクトルではなくてバッファーに最初にロードされた特徴ベクトルであってもよく、そのときのバッファーの内容の全部を正規化係数の計算に利用することができる。この応用例では、遅延の長さはＮであり、Ｎは音声認識の始めでのセグメントの長さである（例えば、Ｎ＝４５）。

【００１６】本発明の１実施例では、特徴ベクトルの成分の全てが正規化されるのではなくて、特徴ベクトルの成分のうちの一部分に対してだけ正規化が実行される。例えば、人の聴覚作用／音声認識に関して最も重要な成分だけに対して正規化を実行してもよい。また、本発明の変形例では、平均値又は標準偏差と関連させて特徴ベクトルに対して正規化を実行するだけでもよい。より一般的に、特徴ベクトルの修正を如何なる統計量に関連させて実行してもよい。

【００１７】図７は移動局の構造を示しており、この移動局には、本発明を利用する音声認識装置６６が設けられている。この移動局は、該装置に特有の例えばマイクロホン６１，キーボード６２，ディスプレイ６３，スピーカー６４及び制御ブロック６５などの部分からなっており、この制御ブロックは該移動局の動作を制御する。また、この図は、移動局に特有の送信ブロック６７及び受信ブロック６８も示している。制御ブロック６５は、該移動局と関連している音声認識装置６６の動作も制御する。この音声認識装置がそのトレーニング段階又は実際の音声認識プロセス時に活性化されているとき、ユーザーが与えるオーディオコマンドが制御ブロックによって制御されてマイクロホン６１から音声認識装置６６に送られる。オーディオコマンドは別のＨＦ（hands free（ハンドフリー））マイクロホンを通して送られてもよい。通常、音声認識装置はＤＳＰによって実現され、その動作に必要なＲＯＭ／ＲＡＭメモリー回路を有する。

【００１８】表２は本発明の方法での音声認識装置の性能を他の騒音補償方法と比べて示している。本発明は、正規化されていないメル周波数セプストラル係数又はＰＭＣ（Parallel Model Combination（並列モデル結合））法の使用と比較されている。試験は、騒音の少ない環境でモデル化されたヒドゥン・マルコフ・モデルを使って実行された。音声認識時には、必要な信号対雑音比を達成するために、認識されるべき単語に雑音信号が加えられた。”クリーン”モードは、音声認識装置のトレーニングと実際の音声認識プロセスとがともに騒音の少ない環境で行われた事態に相当する。試験結果は、本発明の音声認識装置が特に騒々しい環境で認識装置の信頼性を向上させることを証明している。また、本発明の音声認識装置は、計算に関しては本発明の方法よりはるかに複雑なＰＭＣ法より良好に機能することが分かる。

【表２】

環境（SNR）	MFCC- 係数	PMC	正規化された特徴ベクトル
クリーン	96.5%	96.6%	97.5%
5dB	95.0%	95.3%	96.1%
0dB	93.7%	94.9%	95.9%
-5dB	89.3%	93.0%	95.3%
-10dB	73.8%	84.6%	94.3%

【００１９】本明細書では本発明を具体例により説明している。例えば、上の解説では、ＨＭＭに基づく音声認

識装置で本発明を解説している。しかし、本発明は他の手法に基づく音声認識装置に用いるのにも適している。例えば、ニューラル・ネットワークを利用する音声認識装置に本発明を適用することができる。本発明は上記の実施例の詳細に限定されるものではなく、本発明の特徴から逸脱せずに本発明を他の形でも実施し得ることは当業者にとっては明らかなことである。上記実施例は、限定をするものではなくて実例であると解されるべきものである。従って、本発明を実施し使用する可能性は特許請求の範囲の各請求項のみによって限定される。従って、各請求項により確定される、均等実施態様を含む、本発明のいろいろな実施態様も本発明の範囲内に属する。

【図面の簡単な説明】

【図１】従来技術の音声認識装置の構造を示すブロック図である。

【図２】従来技術による分析ブロックの構造を示すブロック図である。

【図３】（Ａ）及び（Ｂ）は、本発明の音声認識装置の構造を示す図である。

【図４】本発明による正規化バッファーの使用を示す図である。

【図５】本発明による方法の作用を示すフローチャート（その１）である。

【図６】本発明による方法の作用を示すフローチャート（その２）である。

【図７】本発明の移動局の構造を示す図である。

【符号の説明】

１１，３０…フロント・エンド

１３…音声認識装置のトレーニングブロック

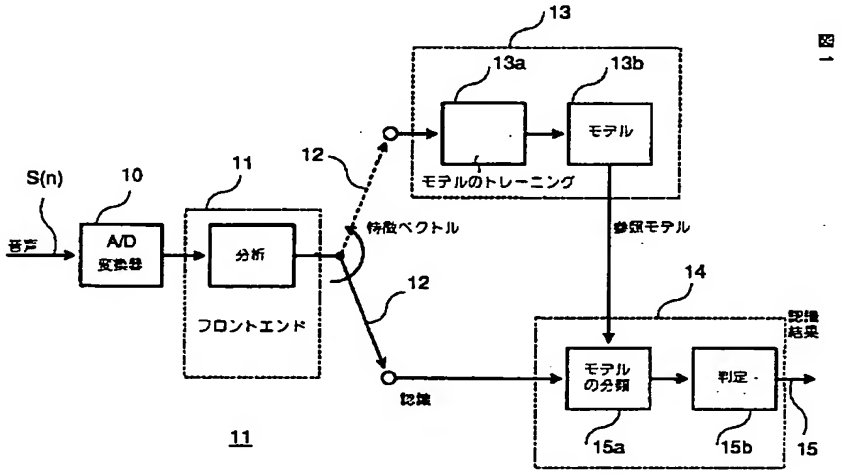
１４…現実の音声認識装置

２０…プリエンファシス・フィルター

２１…フレーム形成ブロック

３１…正規化バッファー

【図１】



【図2】

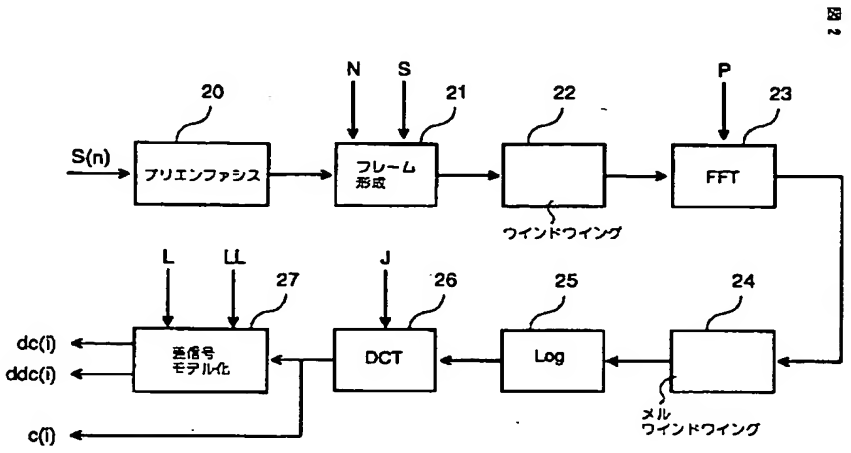


図2
S(n)
20
プリエンファシス
N
S
21
フレーム形成
22
ウインドウイング
P
23
FFT
L
LL
27
J
26
25
24
dc(i)
ddc(i)
c(i)
音響モデル化
DCT
Log
メルウインドウイング

【図3】

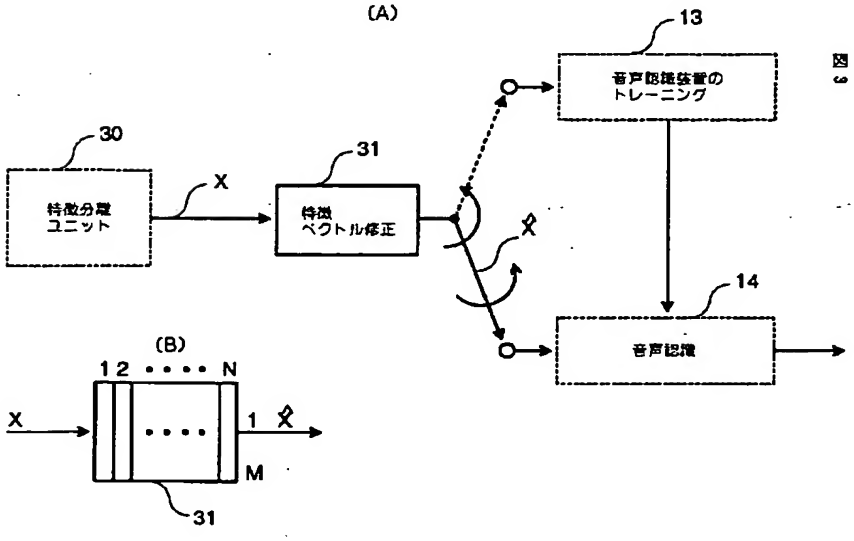
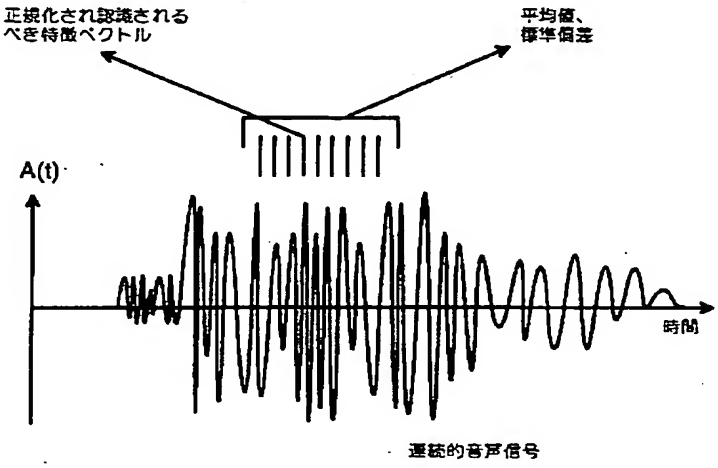


図3
(A)
13
音声認識装置のトレーニング
30
特徴分離ユニット
X
31
特徴ベクトル修正
14
音声認識
(B)
1 2 ・・・・ N
X
1
M
31

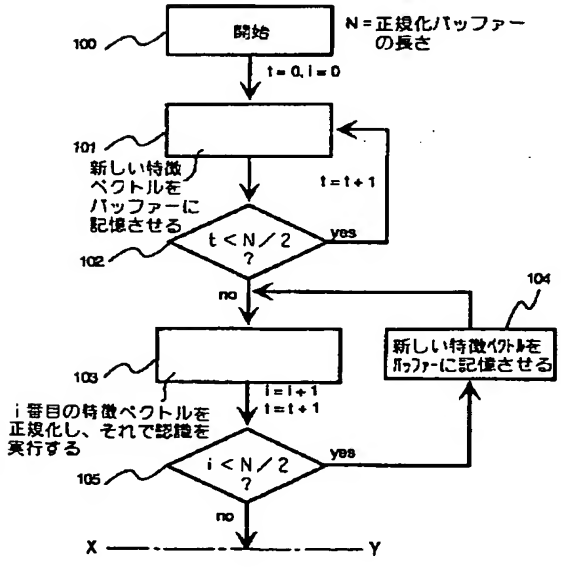
【図4】

図 4



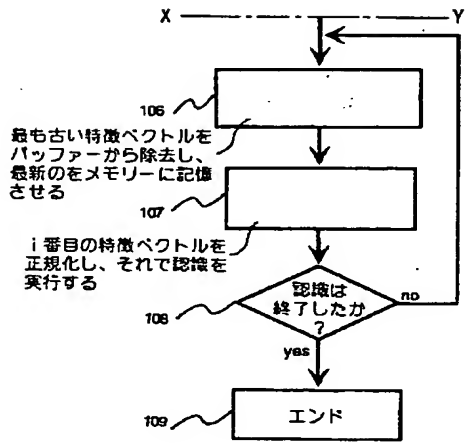
【図5】

図 5



【図6】

図 6



【図7】

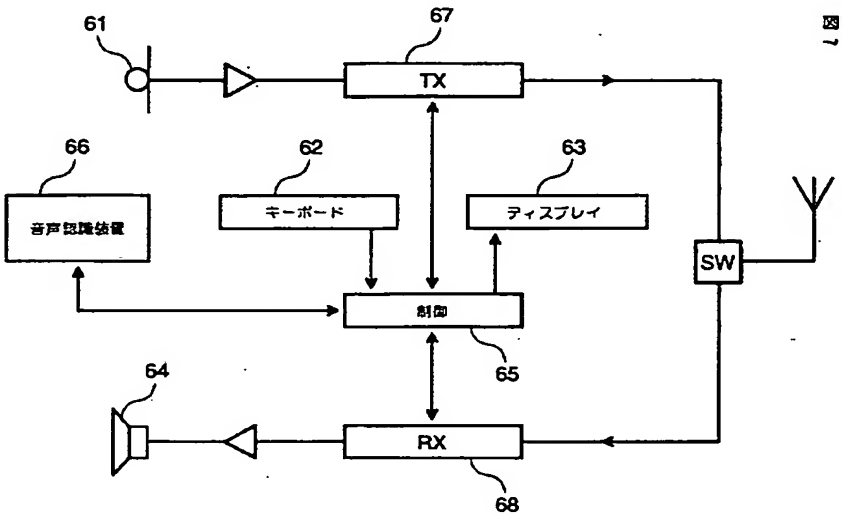


図7